Dynamics of immature mAb glycoform secretion during CHO cell culture

An integrated modelling framework

Jimenez del Val, Ioscani; Fan, Yuzhou; Weilguny, Dietmar

Published in:
Biotechnology Journal

Link to article, DOI:
10.1002/biot.201400663

Publication date:
2016

Document Version
Peer reviewed version

Research Article

Dynamics of immature mAb glycoform secretion during CHO cell culture: An integrated modelling framework

Ioscani Jimenez del Val[1]

Yuzhou Fan[2, 3]

Dietmar Weilguny[3]

[1] School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, Ireland.

[2] Network Engineering of Eukaryotic Cell Factories, Department of Systems Biology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

[3] Symphogen A/S, Pederstrupvej 93, 2750 Ballerup, Denmark.

Correspondence: Dr. Ioscani Jimenez del Val, School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, Ireland.

Email: ioscani.jimenezdelval@ucd.ie

Keywords: Dynamic glycosylation model, *in silico* glycoengineering, CHO cells, pharmaceutical bioprocessing, therapeutic protein glycosylation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/biot.201400663.

Submitted: 11-May-2015
Revised: 09-Oct-2015
Accepted: 23-Dec-2015

Abbreviations: CS, coupling strategy; **Man9,** nine mannose glycan; **Man5,** five mannose glycan; **GE,** glycosylation enzymes; **NS,** nucleotide sugars; $\boldsymbol{q_p}$**,** cell specific productivity; **CD,** cell culture dynamics model; **GlycoModel,** glycosylation model; **GA,** Golgi apparatus; **ER,** endoplasmic reticulum; **ERGIC,** ER-Golgi intermediate compartment; **COPI,** coat protein complex I; **COPII,** coat protein complex II; **GRP,** Golgi resident protein; **ManI,** α-1,2 mannosidase I (EC 3.2.1.113); **GnTI,** α-1,3 N-acetylglucosaminyl transferase I (EC 2.4.1.101); **ManII,** α-1,3/ α-1,6 mannosidase II (EC 3.2.1.113); **GnTII,** α-1,6 N-acetylglucosaminyl transferase II (EC 2.4.1.143); **FucT,** α-1,6 fucosyltransferase (EC 2.4.1.68); **GalT,** β-1,4 galactosyltransferase (EC 2.4.1.38); **SiaT,** α-1,6 sialyltransferase (EC 2.4.99.1)

Abstract

Ensuring consistent glycosylation-associated quality of therapeutic monoclonal antibodies (mAbs) has become a priority in pharmaceutical bioprocessing given that the distribution and composition of the carbohydrates (glycans) bound to these molecules determines their therapeutic efficacy and immunogenicity. However, the interaction between bioprocess conditions, cellular metabolism and the intracellular process of glycosylation remains to be fully understood. To gain further insight into these interactions, we present a novel integrated modelling platform that links dynamic variations in mAb glycosylation with cellular secretory capacity. Two alternative mechanistic representations of how mAb specific productivity (q_p) influences glycosylation are compared. In the first, mAb glycosylation is modulated by the linear velocity with which secretory cargo traverses the Golgi apparatus. In the second, glycosylation is influenced by variations in Golgi volume. Within our modelling framework, both mechanisms accurately reproduce experimentally-observed dynamic changes in mAb glycosylation. In addition, an optimisation-based strategy has been developed to estimate the concentration of glycosylation enzymes required to minimise mAb glycoform variability. Our results suggest that the availability of glycosylation machinery relative to cellular secretory capacity may play a crucial role in mAb glycosylation. In the future, the modelling framework presented here may aid in selecting and engineering cell lines that ensure consistent mAb glycosylation.

1. Introduction

Monoclonal antibodies (mAbs) have emerged as the highest-selling products of the pharmaceutical industry [1]. This unprecedented commercial success can be attributed to the efficacy with which this class of protein therapeutics treat a multitude of life-threatening illnesses, including several types of cancer and autoimmune disorders [2]. All commercially-available mAbs contain two consensus asparagine-linked (N-linked) glycosylation sites on the Cγ2 domains of their constant fragment (Fc) [3]. The monosaccharide composition and distribution of the carbohydrates (glycans) bound to these sites has been widely reported to

influence the safety [4] and therapeutic efficacy of mAbs [3]. In turn, variations in bioprocess conditions have been shown to heavily impact the distribution of glycans bound to therapeutic glycoproteins [5]. Mathematical modelling, in turn, has emerged as an instrument that can provide further qualitative and quantitative insight into the mechanisms by which bioprocess conditions influence therapeutic protein glycosylation [6, 7].

Until recently, dynamic data for mAb glycoform distribution throughout cell culture has been sparingly available in the literature. However, with the development of robust, high-throughput methods for mAb glycan profiling [8-10], a solid knowledgebase on mAb glycosylation dynamics is being generated [9-13]. Interestingly, the majority of these reports present similar qualitative trends, where mAbs with less processed glycans accumulate towards the end of culture. In particular, the changes that have most commonly been reported involve reduced galactosylation [9, 11] and increased secretion of high-mannose glycans (in particular, the Man5 glycan) [10, 13]. These reports highlight the dynamic nature of mAb glycosylation during cell culture and open the possibility of utilising dynamic glycan distribution data to further understand the interplay between cell culture conditions, cell line characteristics and therapeutic protein glycosylation.

Increased secretion of under-processed mAb glycoforms during cell culture is particularly relevant in the context of pharmaceutical bioprocessing. The Man5 glycoform has been associated with decreased serum half-life [14, 15] and compromised effector function activation of mAbs [3, 16]. Furthermore, antibodies bearing this glycan have been associated with detrimental inflammatory responses in humans [17]. Despite these findings, several research groups are exploring the manufacture of Man5 mAb glycoforms to reduce product glycan variability and capitalise on increased antibody dependent cellular cytotoxicity (ADCC) which may result from the absence of core fucose on Man5 glycoforms [13]. Reduced galactosylation, in turn, has been associated with decreased complement activation [18]. Given the influence immature glycoforms have on the therapeutic efficacy of mAbs, understanding the dynamics of

their secretion during cell culture is highly relevant towards developing next-generation therapeutic glycoprotein manufacturing processes.

As reviewed recently [6], several mathematical models have been developed to represent the protein glycosylation process [19-24]. Of these, the models formulated by Umaña and Bailey [24], Krambeck and Betenbaugh [19], Hossler et al. [20] and del Val et al. [21] are single-cell models where glycosylation is assumed to be independent of extracellular conditions, cell growth and recombinant protein productivity. The remaining models have built upon their predecessors by coupling the glycosylation process with cellular metabolism and bioreactor culture dynamics. The work by Ohadi and collaborators [22] couples the glycosylation model of Hossler et al. [20] with a model for CHO cell growth and metabolism and an abbreviated model for nucleotide and nucleotide sugar (NS) biosynthesis. Similarly, the model by Jedrzejewski et al. uses the glycosylation model of del Val et al. [21] and integrates it into a framework for cell culture dynamics and a detailed dynamic model for NS metabolism [23]. In essence, both coupled models sought to describe limitations to mAb glycosylation in terms of the dependency between extracellular nutrient availability and intracellular NS concentrations, and both have been shown to reproduce experimentally-observed glycan distributions accurately. Because both models create a direct link between readily-measurable bioprocess conditions and recombinant product glycosylation, they have been identified as the most amenable for *in silico* glycoengineering in the context of pharmaceutical bioprocessing [6].

Despite successfully coupling cellular metabolism with mAb glycosylation, both of these integrated models do not specify whether the rate of recombinant protein secretion (q_p) was considered to influence protein glycosylation. Indeed, both these processes (protein secretion and glycosylation) are mediated by Golgi apparatus (GA) dynamics [25, 26] and, thus, are inextricably linked to one another. The relationship between GA dynamics, q_p and glycosylation is presented in Figure 1. Crucially, the mechanisms described in Figure 1 define the rate at which secretory cargo enters (and leaves) the Golgi space, and how these dynamics influence

the morphology and resident protein content of the GA. In order to mathematically model the interplay between q_p and protein glycosylation, two alternative representations of Golgi dynamics are considered in this work:

i. CS1: Increased q_p (compared to baseline levels – Figure 1A) may occur due to secretory cargo traversing the Golgi apparatus (GA) at higher rates [27, 28]. This would reduce the retention time of glycoproteins within the GA, and thus, decrease the extent of glycosylation reactions therein [20] (Figure 1B).

ii. CS2: Increased q_p could also arise from an enlarged secretory (Golgi) apparatus [29, 30]. In this case, if glycosylation enzyme (GE) expression levels were to remain unchanged, their availability within a larger GA would be diluted (Figure 1C). In turn, lower GE concentrations would have a negative impact on the rate of glycosylation reactions.

Interestingly, most studies where increased q_p has been induced (e.g. by reducing temperature [31-34] or by adding supplements such as sodium butyrate [34-37]) report variable consequences on glycosylation and have generally attributed them to pleiotropic effects associated with cell cycle arrest. However, several publications have linked increased secretory capacity (q_p) with reduced glycoform complexity [32, 38, 39]. Of these, the recent publication by Fan et al. [39] integrates experimental data for intracellular NS availability, mAb glycoprofiling and comparative and quantitative proteomics to conclude that a likely cause for increased secretion of immature glycoforms (particularly hypogalactosylated and Man5 glycans) is reduced glycosylation machinery relative to cellular secretory capacity.

In light of the possible correlation between q_p and glycan processing, we propose a novel modelling framework where the influence of recombinant protein secretion rate is mechanistically linked to the glycosylation process within the GA. The two interpretations of how increased secretory capacity influences the glycosylation process within Golgi (described as *i* and *ii*, above and presented in Figure 1) are compared in this work. Using the experimental

data reported by Fan et al. [39], we present a comprehensive strategy to estimate the unknown parameters of our model. Once parameterised, the model quantitatively reproduces the experimentally-observed increase in immature mAb glycoform secretion at higher rates of product secretion. Results from our modelling framework suggest that availability of glycosylation machinery relative to mAb specific productivity may indeed be responsible for increased secretion of under-processed glycans during CHO cell culture. Additional optimisation-based calculations within our modelling framework indicate that a ~3-fold increase in the expression of functional N-acetylglucosaminyltransferase I (GnTI – the enzyme responsible for the transition from high mannose to hybrid and complex glycans) should maintain secretion of Man5 glycans at baseline levels. The modelling strategy presented in this work provides additional insight into cellular processes which may negatively impact glycan processing. Thus, it has potential applications in refining selection and glycoengineering strategies to identify and generate cell lines that are capable of producing consistent and optimal mAb glycoforms.

2. Mathematical model development

2.1. Experimental data calculations

Three distinct phases of culture were observed in the experimental data used for this work [39]. Phase I is characterised by culture at 37°C and occurs between 0 and 5 days (0 to 120h). Phase II is defined as culture at 33.5°C and spans from days 5 to 9 (120 to 216h). Finally, Phase III is culture at 33.5°C with stationary cell growth and was observed from day 9 (216h) until the end of culture. Estimates for average values of net specific growth rate (μ_{net}), specific glucose uptake rate (q_{glc}) and specific mAb secretion (q_p) were obtained from the experimental data by calculating the slope of the cumulative production/uptake of each species with respect to the integral viable cell number [31]. Details on these calculations are presented in Supplementary Section 1.

2.2. Integrated mathematical model for mAb glycosylation

An integrated model to describe the dynamics of mAb glycosylation throughout cell culture was formulated by coupling an unsegregated and unstructured cell culture dynamics (CD) model with the model for intracellular mAb glycosylation (GlycoModel) described by del Val et al. [21]. A schematic representation of the modelling strategy employed in this work is presented in Figure 2. The assumptions underlying all models integrated in this work are presented in Supplementary Section 2. All simulations, optimisations and parameter estimations were performed with gPROMS ModelBuilder 4.1.0 (www.psenterprise.com/gproms) [40].

2.2.1. Cell culture dynamics model

An unsegregated and unstructured cell culture dynamics model was developed to represent cell growth and death, glucose uptake, lactate secretion and uptake, as well as mAb production. The CD model material balances were defined as Equations 1 through 6.

$$\frac{dV}{dt} = Q_{feed} - Q_{out} \quad \text{Eq. 1}$$

$$\frac{d(V[X_v])}{dt} = -Q_{out}[X_v] + (\mu_g - k_d)[X_v]V \quad \text{Eq. 2}$$

$$\frac{d(V[X_d])}{dt} = -Q_{out}[X_d] + k_d[X_v]V \quad \text{Eq. 3}$$

$$\frac{d(V[Glc])}{dt} = Q_{feed}[Glc]_{in} - Q_{out}[Glc] - q_{glc}[X_v]V \quad \text{Eq. 4}$$

$$\frac{d(V[Lac])}{dt} = -Q_{out}[Lac] + (q_{lac,secr} - q_{lac,cons})[X_v]V \quad \text{Eq. 5}$$

$$\frac{d(V[mAb])}{dt} = -Q_{out}[mAb] + q_p[X_v]V \quad \text{Eq. 6}$$

The specific rates for cell growth (μ_g) [41], cell death (k_d) [42], glucose uptake (q_{glc}), lactate consumption ($q_{lac,cons}$), lactate secretion ($q_{lac,secr}$) and mAb specific productivity (q_p) were defined as Eqs. 7–12.

$$\mu_g = \mu_{g,max}\left[\left(\frac{[Glc]}{K_{m,glc}+[Glc]}\right)-\frac{[X_v]}{\alpha_X}\right] \quad \text{Eq. 7}$$

$$k_d = k_{d,max}\left(\frac{K_{d,\mu}}{K_{d,\mu}+\mu_g}\right) \quad \text{Eq. 8}$$

$$q_{glc} = \left(\frac{1}{Y_{X/glc}}\right)\left(\frac{[Glc]}{K_{m,glc}+[Glc]}\right) \quad \text{Eq. 9}$$

$$q_{lac,secr} = \left(Y_{lac/glc}\right)q_{glc} \quad \text{Eq. 10}$$

$$q_{lac,cons} = \left(\frac{1}{Y_{X/lac}}\right)\left(\frac{[Lac]}{K_{m,lac}+[Lac]}\right) \quad \text{Eq. 11}$$

$$q_p = \left(Y_{mAb/glc}\right)q_{glc} \quad \text{Eq. 12}$$

The fundamental criterion for development of the CD component was to generate the simplest model possible that could capture cell growth and mAb productivity accurately, given that these variables are the key inputs for the GlycoModel. Details on the assumptions underlying the CD model are presented in Supplementary Section 2.

The CD model contains the ten kinetic parameters listed in Supplementary Table ST1. Given the distinct behaviour of cell culture at 37°C and 33.5°C, two sets of parameters, one for each temperature, were defined [31]. Parameter values for the CD model were obtained using the maximum likelihood parameter estimation activity available in gPROMS 4.1.0, where the values for the unknown physical and variance model parameters are obtained so that the probability of the model predicting the experimental values is maximised [40]. Data for viable cell density, dead cell density, extracellular glucose, lactate and product were used for parameter estimation.

2.2.2. Glycosylation model (GlycoModel)

The three equations defining the main features of the glycosylation model used in this work are presented as Eq. 13–Eq. 15 [21].

$$\frac{\partial[Glyc_i]}{\partial t} = -V_l \frac{\partial[Glyc_i]}{\partial z} - \sum_{j=1}^{N.R.} \nu_{i,j} r_j \qquad \text{Eq. 13}$$

$$[E_j](z) = E_{j,max} exp\left[-\frac{1}{2}\left(\frac{z - z_{j,max}}{\omega_j}\right)\right]^2 \qquad \text{Eq. 14}$$

$$r_j = \frac{k_{f,j}[E_j][NS_k][Glyc_i]}{K_{d,i/j}K_{d,k/j}\left(1 + \frac{[NS_k]}{K_{d,k/j}} + \frac{[NS_k]}{K_{d,k/j}}\frac{[Glyc_i]}{K_{d,i/j}} + \frac{[NS_k]}{K_{d,k/j}}\sum_{n=1}^{N.C.}\frac{[Glyc_n]}{K_{d,n/j}} + \frac{[B_k]}{K_{i,k/j}}\frac{[Glyc_{i+1}]}{K_{d,(i+1)/j}} + \frac{[B_k]}{K_{i,k/j}}\right)} \qquad \text{Eq. 15}$$

Eq. 13 represents the material balances of maturing cisternae flowing through a Golgi space of length z (Figure 2B). There, the inlet of secretory cargo into the Golgi is determined by the concentration of glycan i ($Glyc_i$) at $z = 0$ multiplied by V_l, which represents the linear velocity with which the secretory cargo traverses the Golgi space. GRP recycling is represented by defining the distribution of these species along the Golgi space length as three-parameter normal functions (Eq. 14), where $E_{j,max}$ represents the peak enzyme concentration, $z_{j,max}$ the localization of the peak concentration and ω_j, the breadth of enzyme concentration. Eq. 15 is presented as an example of the rate expressions included in the GlycoModel. The reaction network considered in this model was developed to represent mAb Fc glycosylation and consists of 77 glycans linked through 95 reactions catalysed by eight enzymes [21].

Given that intracellular availability of NSs was not reported to control secretion of immature glycans in the dataset used for this work [39], their concentration in the GA was assumed to remain constant throughout culture. Golgi concentration of UDP-GlcNAc and UDP-Gal was derived from the day 5 (120h) values reported for this dataset. These data were selected because they were the lowest observed concentrations for both NSs during culture. Even at these low abundances, the respective enzymes (GnTI, GnTII and GalT) would be saturated with

NSs according to the literature values for $K_{d,k}$ (dissociation constants of NSs from the enzymes) [43-45]. All considered values for Golgi NS concentrations, along with their derivation are presented in Supplementary Section 3 and Table ST2.

2.2.3. Model coupling strategy

In contrast to previous integrated models, the work presented herein defines a direct mechanistic link between mAb q_p and the intracellular process of glycosylation (Eq. 16). In addition, Eq. 17 has been defined to represent the accumulation of mAb glycoform i (mAb_i) in the extracellular environment as a function of the glycan molar fraction at the exit of Golgi, $\left[x_{mAb,i}\right]_{z=1}$ [22, 23].

$$(V_l)\left(Vol_{Golgi}\right)[Man9]_{z=0} = \left(\frac{q_p}{\overline{MW_{mAb}}}\right)\left(2\frac{\mu mol_{Glyc}}{\mu mol_{mAb}}\right) \quad \text{Eq. 16}$$

$$\frac{d(V[mAb]_i)}{dt} = Q_{out}[mAb]_i + \left[x_{mAb,i}\right]_{z=1} q_p[X_v]V \quad \text{Eq. 17}$$

In Eq. 16, the linear velocity with which secretory cargo traverses the Golgi space (V_l), Golgi volume (Vol_{Golgi}) and the concentration of Man9 glycoform entering the Golgi space ($[Man9]_{z=0}$) are all proportional to q_p. Given that one of the major functions of the ERGIC is to concentrate secretory cargo prior to its transport to the GA [46], and that the cargo contents at the inlet of the Golgi space are determined by the number of secretory vesicles that fuse to generate new cisterna [25, 30], it is possible to assume that the value for $[Man9]_{z=0}$ remains relatively constant. This parameter's value was obtained with Eq. 16 using a $q_p = 12.3\frac{pg}{cell\ day}$ (the highest experimentally observed [39]), $V_l = 0.045\frac{Golgi\ Length}{min}$ (as derived from the Golgi retention time reported previously [21, 24]), and an average mAb molecular weight ($\overline{MW_{mAb}}$) of 150kDa. This calculation yields a value of $[Man9]_{z=0} = 100\frac{\mu mol}{L_{Golgi}}$. It is worth noting that the highest q_p observed in the considered experimental dataset [39] was used for this calculation to

ensure the obtained value for $[Man9]_{z=0}$ is consistent with those reported previously [19-21, 23, 24].

As discussed in Section 1, V_l and Vol_{Golgi} may vary [28, 29], thus potentially impacting Golgi dynamics and, consequently, glycosylation. With these considerations, two "coupling strategies" have been defined to isolate the influence each of these variables has on the dynamics of mAb Fc glycosylation. In the first coupling strategy, Vol_{Golgi} is assumed to remain constant, while the variable which controls Golgi dynamics is V_l. In the second coupling strategy, V_l is assumed to remain constant while Vol_{Golgi} varies, thus controlling Golgi dynamics.

2.2.3.1. Coupling Strategy 1 (CS1): variable linear velocity through Golgi

Here, the inlet and outlet rates of secretory cargo traversing the Golgi space (the linear velocity through Golgi) are assumed to be equal and proportional to q_p. This assumption is supported by experimental evidence where variable transport rates between the ER and Golgi have been observed in living cells [28]. Moreover, anterograde dynein/dynactin-mediated vesicle transport has also been observed to occur at variable rates [27]. A fundamental consequence of the inlet and outlet rates being equal is that no changes in the volume of Golgi are assumed to occur, regardless of changes in q_p. In addition, we assume that COPI vesicle-mediated retrograde transport of GRPs for recycling occurs at a sufficiently high rate that ensures a quasi-steady state concentration of resident proteins within Golgi when no change in resident protein expression occurs. Having defined the values for $\overline{MW_{mAb}}$, Vol_{Golgi} and $[Man9]_{z=0}$, V_l was made a function of the q_p calculated by the CD model and was linked to the GlycoModel through Eq. 13.

2.2.3.2. Coupling Strategy 2 (CS2): variable Golgi volume

CS2 assumes that the rate at which secretory material leaves the Golgi space is not equal to that at which it arrives. As a consequence, increased transport from ERGIC to the Golgi space results in an accumulation of Golgi structures, which in turn, leads to increased volume of the GA.

Increased product secretion is achieved by a larger volume of secretory components leaving the GA at a constant rate. This behaviour is supported by experimental evidence where the volume of ER and Golgi has been observed to increase upon induced synthesis of a recombinant protein [29] and a detailed study where increased expression of secretory pathway components was observed at very high q_p (>100pg/cell/day), while the secretion rate (V_l) was observed to remain relatively constant [34].

Changes in Golgi volume were implemented by defining the dilution function presented in Eq. 18. There, variations in Vol_{Golgi} are calculated based on the q_p from the CD model using Eq. 16 while assuming constant V_l and $[Man9]_{z=0}$. The baseline Golgi volume ($Vol_{Golgi,baseline}$), was defined by solving Eq. 16 using constant values for $V_l = 0.045 \frac{Golgi\ Length}{min}$ [21, 23] and $[Man9]_{z=0} = 100\mu M$, along with a $q_p = 5.6 \frac{pg}{cell\ day}$, as derived from the experimental data for Phase I of culture. This value of q_p was used, given it was the lowest observed in the experimental dataset considered for this work [39]. The resulting $Vol_{Golgi,baseline}$ was $1.15 \times 10^{-14} L$, which is within the same order of magnitude of Golgi volumes assumed in previous models [19-21, 24].

$$f_{dil} = \frac{Vol_{Golgi}}{Vol_{Golgi,baseline}} \quad \text{Eq. 18}$$

Considering that no large variations in GRP expression (neither enzymes or NS transporters) were reported for our considered dataset [39], the concentrations of glycosylation enzymes (GE) within Golgi (Eq. 14) were multiplied by the dilution coefficient (Eq. 18) to represent q_p-associated variations in Golgi volume. An important additional consideration was that increases in Golgi volume were assumed to be irreversible within the timeframes associated with transient reductions in q_p. The mechanism underlying this assumption is that the redistribution of secretory machinery from Golgi to other components of the pathway occurs over longer periods of time when compared to the short-lived reductions in product secretion.

Despite the assumed changes in Golgi volume, the concentration of NSs and secretory cargo in this organelle are assumed to remain constant. As discussed previously, the reported concentrations of NSs (even when considering maximal Golgi volume dilution) were observed to saturate their respective enzymes. In turn, the concentration of secretory cargo was assumed constant given the dependence between cis-Golgi cisterna formation and the cargo capacity of COPII-coated vesicles. For all simulations with CS2, V_l and $[Man9]_{z=0}$ were assumed to remain constant and equal to $0.045\,\frac{Golgi\ Length}{min}$ and $100\mu M$, respectively.

2.2.4. GlycoModel parameter estimation

Estimation of the unknown parameters in both CSs was performed in a series of sequential steps based on previous reports [21]:

- Step 1: The fraction of mAb glycoform *i* (f_i) secreted during the cell culture interval between $288h$ and $336h$ was calculated with Eq. 19 using the mAb glycan distributions reported at each time point ($Glyc_{i,288h}$ and $Glyc_{i,336h}$, respectively) and the corresponding mAb titres ($[mAb]_{288h}$ and $[mAb]_{336h}$). A full derivation of Eq. 19 is available in Fan et al. [39] and is also presented in Supplementary Section 4 of this work.

$$f_i = \frac{\left(Glyc_{i,t_{n+1}} - Glyc_{i,t_n}\right)}{\left([mAb]_{t_{n+1}} - [mAb]_{t_n}\right)} \qquad \text{Eq. 19}$$

- Step 2: The parameters defining the distribution and concentration of the GnTI, ManII, GnTII and FucT glycosylation enzymes (GEs) along the Golgi space ($E_{j,max}$ and ω_j in Eq. 14) were estimated by obtaining the minimum amount of enzyme necessary to achieve the f_i distribution obtained from Step 1, while also considering a $q_p = 12.3\,\frac{pg}{cell\ day}$. For this estimation, the values for $K_{d,i}$ were kept at the nominal (literature) values shown in Supplementary Table ST3 (Suppl. Section 5). The interval between $288h$ and $336h$ was selected for these estimations because it presented the highest q_p, and thus, excluded

secretion of other high mannose glycans (Man9-Man6) that were not observed experimentally [39].

- Step 3: With the parameters obtained in Step 2, the distribution and concentration of GalT and SiaT were estimated by obtaining the minimum amount of enzyme required to achieve 1% of the G2FS1 (Figure 2) glycoform.
- Step 4: With the distribution of all GEs defined through Steps 2 and 3, three $K_{d,i}$s for GalT (each accounting for the glycan branch specificity reported for this glycosyltransferase [45]) and one for SiaT were estimated using experimentally-derived data for the mAb glycoform distribution determined during Phase I of culture and a corresponding value for q_p of 5.6pg/cell/day (Phase I). This glycoform distribution was used for $K_{d,i}$ estimation because it presented the highest fraction of galactosylated and sialylated glycoforms, thus facilitating accurate and robust estimates for the parameters.

Steps 2 and 3 were performed using the built-in multiple-shooting control vector parametrization algorithm available in gPROMS 4.1.0 (www.psenterprise.com/gproms) [40]. The $K_{d,i}$ values were obtained using the maximum likelihood parameter estimation activity available in the same software package which has been briefly described in the CD model section. Unless stated otherwise, all other parameters for Golgi enzyme localisation ($z_{j,max}$) and enzyme kinetics ($k_{f,j}$, $K_{d,k}$) were considered to be the same as those used by del Val et al. [21] and are presented in Supplementary Tables ST3 and ST5 (Suppl. Section 7). All estimations were performed on stand-alone, steady-state versions of the GlycoModel. We must highlight that the only required data inputs for estimation of all GlycoModel parameters were the mAb glycan distributions secreted during the initial and final phases of culture, along with the corresponding values for q_p for each phase.

3. Results

3.1. Culture dynamics model performance █What are the major results?█

The estimated parameters for the CD model are presented in Supplementary Table ST4 (Suppl. Section 6) and simulation results with the obtained parameters are compared with the experimental data in Figure 3. The ultimate purpose of the CD model was to provide an accurate representation of cell growth and mAb productivity given that these were the critical variables used for coupling with the GlycoModel. As can be seen in Figure 3, these culture variables were captured accurately and show good agreement with the experimental data across all phases of culture.

3.1.1. GlycoModel enzyme distribution estimation█What are the major results?█

The distribution and concentration of glycosylation enzymes (GEs) along the Golgi space were estimated as described in Section 2.2.4. The estimated GE distribution parameters for CS1 and CS2 are presented in Supplementary Table ST5 (Suppl. Section 7) and the resulting profiles along the Golgi space are shown in Figures 4A (CS1) and 4B (CS2). Although the concentrations diverge in magnitude (particularly for ManI, GnTII and FucT), their localisation and distribution are similar to those reported previously [21]. The estimated enzyme profiles show high co-localisation of the ManI and GnTI enzymes, a fairly narrow profile for FucT and relatively low abundances of the GalT and SiaT enzymes for both coupling strategies.

The main differences between the enzyme profiles estimated for both CSs are for ManI, GnTII and FucT. With the exception of GnTII, the profiles are narrower and the peak enzyme concentration is higher for CS1, which is likely due to the assumptions underlying each model. In CS1, the higher linear velocities of secretory cargo through Golgi reduce the time the glycoprotein substrates are exposed to the GEs therein. In order to compensate for this reduced reaction time, higher peak enzyme concentrations were estimated and, because a total enzyme (integral over the Golgi space) minimisation criterion was included, the width of the profiles

was estimated to be narrower. It is possible that this trend was not observed for GnTII because the turnover rate considered for this enzyme is the highest of the glycosyltransferase enzymes included for the reaction network (Supplementary Table ST3). Analogously, the lowest turnover rate corresponds to FucT, and this may be the reason why the profile for this enzyme presents such a high peak concentration. Conversely, the enzyme distributions for CS2 present lower and broader profiles. In this case, the holding time within Golgi remains constant because the linear velocity through Golgi is assumed to remain constant, and thus, lower peak enzyme concentrations are able to achieve the same reaction conversions.

For both coupling strategies, the profiles for GalT and SiaT were found to be similar. This is likely due to the low values for G2F and G2FS1 assumed for their estimation. It is worth noting that, although these profiles provide sufficient activity to satisfy the constraints defined for parameter estimation, availability of these enzymes is likely to be higher *in vivo*. Specifically, the relatively low galactosylation and sialylation of mAb glycans has commonly been attributed to limited access of GalT and SiaT to the glycans present within the crystallisable fragment of mAbs due to steric constraints [47]. Thus, the limited secretion of galactosylated and sialylated glycans would be influenced more heavily by the affinity of the enzymes for the glycans ($K_{d,i}$) than enzyme availability.

3.1.2. GlycoModel $K_{d,i}$ estimation and integrated model simulation results█What are the major results?█

Once the parameters for enzyme concentration and distribution were obtained, the dissociation constants of the glycoproteins from the GEs ($K_{d,i}$) were estimated as described in section 2.2.4. The obtained parameters, along with the respective confidence intervals, are presented in Supplementary Table ST6 (Suppl. Section 8). Simulation results for both integrated models (CS1 and CS2) obtained with the estimated CD, enzyme and $K_{d,i}$ parameters are compared with the experimental data reported by Fan et al. [39] in Figures 4C-4H. There, it can be seen that both

coupling strategies follow very similar trends and reproduce the dynamic glycoform distribution data well.

Of particular relevance to this work is that both coupling strategies are capable of reproducing the shift from a distribution of relatively mature mAb glycoforms during the initial phase of culture to one with considerably higher proportions of Man5 and hypogalactosylated glycans towards the end of culture. Specifically, both coupling strategies predict an increase in the fraction of Man5 glycoform from 2.9% at 120h of culture to 21%, at 336h (Figure 4C). Similarly, the fractions of G1F glycoform (both on the α-1,6 and α-1,3 arms of the biantennary glycan structure) were predicted to drop from 20% to 10% (Figure 4F) and from 9% to 4%, respectively (Figure 4G).

The largest deviations between model and experiment are observed for the relative amount of Man5 present observed during phase III of culture (Figure 4A). Both coupling strategies underestimate the fraction of Man5 at 288h by approximately 2.2%. Additionally, smaller deviations can be seen for the fraction of galactosylated glycans during Phase III of culture. Both coupling strategies overestimate galactosylation by between 1% and 2% during the late phases of culture. If the precipitous drop in glycan complexity indeed occurs due to increased q_p, then underestimation of Man5 and overestimation of galactosylation towards the end of culture may possibly be due to external effects that have not been considered in the model. Such external effects could be depletion of Mn^{2+} (which is a known cofactor for both GnTI and GalT [10, 48]), or inhibition of enzyme activity due to accumulation of metabolites, such as ammonia or lactate [49], towards the latter stages of culture.

3.1.3. Man5 secretion minimisation█What are the major results?█

Considering that a likely cause for the observed increased secretion of Man5 is insufficient availability of GnTI relative to increased secretory capacity [39], both CS models were used to determine the degree of GnTI overexpression that would keep Man5 secretion to its baseline levels (prior to the temperature shift). In order to do so, an optimisation function was defined to

obtain the amount of GnTI required to maintain Man5 secretion below 2% (baseline level). While only estimating $E_{GnTI,max}$, an inherent assumption in this strategy is that all overexpressed GnTI localises in the same region of Golgi.

The resulting enzyme profiles for each coupling strategy are compared with the original ones in Figure 5A and 5C. There, the increase in GnTI concentration is apparent, although the localisation and width of the estimated profiles are similar to those obtained prior to enzyme overexpression. Through this strategy, the obtained $E_{GnTI,max}$ values were found to be $0.93\mu M$ for CS1 and $0.84\mu M$ for CS2 (Supplementary Section 9). With these values, the total concentrations of GnTI within the Golgi space (the integral of $[GnTI]$ over the length of Golgi) were calculated to be $1.82\mu M$ (CS1) and $1.98\mu M$ (CS2), thus representing a 3-fold increase in GnTI concentration compared to the original (high Man5 secretion) values for each coupling strategy.

The simulated end-point glycoform distributions for both coupling strategies after *in silico* GnTI overexpression are presented in Figures 5B and 5C. There, it can be seen that Man5 is maintained at the desired low levels and that reduced secretion of this glycoform results in higher accumulation of more processed glycans, in particular G0F and G1F. Specifically, Man5 is predicted to decrease from 21% to 1.6% (CS1) and 21% to 1.5% (CS2); G0F increases from 56% to 69% for both coupling strategies; and the fraction of G1F goes from 14.3% (CS1) and 14.7% (CS2) to 17.2% and 17.8%, respectively. Dynamic profiles obtained with *in silico* GnTI overexpression are presented in Supplementary Figures S3 and S4.

These results highlight the *in silico* glycoengineering capabilities of our integrated modelling framework and simultaneously provide additional evidence in support of a scheme where the availability of glycosylation machinery relative to cellular secretory capacity influences the glycosylation process.

4. Discussion

4.1. Model formulation █What are the major results and conclusions?█

Based on previous experimental work, our motivation to develop the model presented here was to provide mechanistic links between variations in q_p and the glycosylation process so that dynamic changes in mAb glycoform secretion during CHO cell culture could be described. A key property of the experimental data used for this work was that expression of glycosylation machinery was not observed to vary at different specific productivities [39]. This allowed us to target the abundance of glycosylation machinery relative to cellular secretory capacity as the likely mechanism causing decreased glycan processing. In doing so, an obvious alternative – that glycosylation enzyme activity may be compromised by low temperature culture – was not considered. This phenomenon was not examined given that, to our knowledge, there are no reports associating reduced GnTI activity with lower temperatures. Indeed, animal and plant GnTI has been reported to have a broad range of temperature optima spanning from 25°C to 50°C [50], although studies cited therein refer to recombinant versions the enzyme produced in *E. coli*.

In parallel to this, GnTI activity has been reported to be negatively affected by accumulation of ammonia during cell culture in a pH(Lactate)-dependent manner [49]. We speculate that deviations between model and experiment observed during Phase III of culture may be due to these effects, which were not included in our modelling framework. An additional limitation to our model is that, because NS availability is assumed to remain unchanged, a direct link with cellular metabolism is not provided.

However, the integrated modelling strategy presented here is modular. Thus, the CD model can be exchanged for more detailed descriptions of culture dynamics that include cellular metabolism and cell population dynamics [51, 52]. More specifically, detailed CD models that include a detailed description of NS metabolism [23] could be readily coupled to the GlycoModel via q_p and the NS transport functions (from the cytosol to Golgi) that are already available in the

GlycoModel [21]. We must highlight that, despite not including additional metabolic effects, the major shift to Man5 and hypogalactosylated glycan secretion (reported between Phases I and II [39]) was reproduced accurately by both integrated models, indicating that q_p may indeed impact recombinant protein glycosylation.

In developing the two coupling strategies, we aimed to identify which process was most likely to generate the variations in mAb glycoform secretion observed in the experimental data. As discussed extensively throughout the manuscript, CS1 assumes that q_p determines the linear rate of secretory cargo through the GA. In support of this assumption is previous work where Golgi dynamics have been visualised in living cells [28] and where the rate at which dynein-dynactin mediated transport of secretory vesicles has been found to be variable [27]. Furthermore, a mechanism whereby increased expression of GEs leads to higher COPII assembly and, consequently to increased Golgi volume has been observed experimentally [30]. The proteomic data obtained for the considered experimental dataset found no significant increase in GEs or proteins involved in COPII-vesicle formation at different q_ps [39], suggesting that *de novo* synthesis of Golgi components may not have been prevalent during culture.

The above findings contrast with previous work that has observed increased expression of COPI and COPII proteins in cell lines that present higher specific productivity [53]. Increased expression of COPII proteins, in particular, could result in higher *de novo* synthesis of Golgi components and, possibly, lead to increased Golgi volume [30], thus favouring the mechanism depicted by CS2. In addition, secretion kinetics have been reported to remain relatively constant while an overall increase in secretory pathway components was observed during culture of hyper-productive CHO cells treated with sodium butyrate under mild hypothermic conditions [34]. In line with these findings, increases in Golgi volume during synthesis of recombinant proteins has been observed [29]. However, while using the same methods for determining Golgi volume, recent reports have found no correlation between the size of Golgi and increased specific productivity [54]. In conjunction, the above seemingly conflicting findings suggest that

both mechanisms (variable V_l or variable V_{Golgi}) could occur simultaneously and prevalence of one over the other may be cell line dependent.

4.2. Integrated model performance █What are the major results and conclusions?█

In terms of reproducing the dynamic changes in glycoform distribution during culture, both CSs perform similarly well. The reason behind this is that they both present two alternative mechanisms that result in the same phenomenon. In both cases, impaired glycan processing with increased q_p is defined to occur due to reduced exposure of secretory cargo to the GEs within Golgi. In CS1, reduced exposure occurs due to lower Golgi holding times associated with variations in V_l, while CS2 represents reduced exposure by considering that the GEs are diluted through a q_p-induced increase in Golgi volume. Given the similarities in model performance, the current framework cannot identify which mechanism controls the process for this particular cell line.

Regardless of the controlling mechanism, we must highlight that adequate model performance supports the hypothesis that availability of glycosylation machinery relative to cell specific productivity is the underlying cause behind secretion of less processed mAb glycoforms upon a shift to increased q_p. In particular, estimation of the glycosylation enzyme distribution parameters was performed using a steady-state version of the GlycoModel utilising a single data-point (corresponding to the time interval between 288h and 336h). With the obtained parameters, the changes in q_p described by the CD model (which was parameterised independently) generated the experimentally-observed shift to increased Man5 secretion without any additional input. This was observed for both coupling strategies, and in itself, indicates that the relationship between q_p and availability of glycosylation machinery interact to *quantitatively* reproduce the dynamic changes observed in mAb glycoform secretion.

With these results and previous experimental evidence, it is possible to conclude that the relative abundance of glycosylation machinery relative to cellular secretory capacity may play a key role in the glycosylation process. To our knowledge, this has not been confirmed experimentally, although specific productivity has been identified to potentially influence recombinant protein glycosylation [32, 38, 39]. Contrastingly, several reports where product glycosylation has been monitored under high productivity conditions (temperature shift and sodium butyrate treatment) have not found significant impacts on product glycosylation [33, 37].

Taking these findings into consideration and linking them with the results of our modelling framework, it is possible that the correlation between q_p and reduced glycan processing has not been more broadly established due to cell line-specific variations. A shift to secretion of less processed glycans would only be observed if glycosylation machinery is sufficiently available at low q_ps, but becomes limiting at higher productivities. If the baseline levels of glycosylation machinery are high enough to cope with higher q_ps, the distribution of secreted glycans would remain relatively unchanged. It would be interesting to see if these properties are observed in a cell line that exhibits very similar dynamic trends in Man5 secretion during culture [10].

Industrially, the ramifications of this hypothesis would extend to cell-line selection and bioprocess optimisation. A cell line with relatively low baseline levels of, for example, GnTI may be selected for further development, given that, under the examined conditions, it produces acceptable levels of Man5. However, when cultured under conditions that increase productivity, an undesired shift to higher Man5 glycoform secretion may occur due to limitations in GnTI availability relative to a higher q_p. From this perspective, monitoring expression of different GEs relative to specific productivity during early stages of bioprocess development may prove useful in reducing the probability of encountering shifts in glycosylation-associated quality during latter stages of bioprocess development.

Results obtained with the *in silico* GnTI overexpression strategy described in this work provide additional support for the hypothesis that correlates variations in q_p with decreased secreted glycan complexity. While simulating overexpression of GnTI, enough of this enzyme is provided to maintain a high enough rate of Man5 processing even at the highest q_p values. Thus, secretion of this immature mAb glycoform is minimised. Although the results obtained from *in silico* GnTI overexpression are yet to be experimentally confirmed, such a computational strategy could be used to estimate the expression levels of GEs required to maintain certain product glycoform distribution at predefined levels of productivity. This, in turn, could aid in cell-line selection and engineering during early stages of bioprocess development.

In conclusion, the modelling framework presented herein is the first to provide a detailed mechanistic link describing how cell specific productivity may impact monoclonal antibody glycosylation. This computational tool could aid in cell line selection and engineering during the early stages of bioprocess development, and if linked with more detailed models for cellular metabolism, could be employed to control pharmaceutical bioprocesses to ensure production of therapeutic proteins with consistent and optimal glycosylation profiles.

Acknowledgments

We would like to thank Dr Mikael Rørdam Andersen, Prof Michael J Betenbaugh and Dr Cleo Kontoravdi for providing the experimental data used for this work. The gPROMS code for the GlycoModel was kindly provided by Dr Cleo Kontoravdi. IJV would also like to thank Prof Mohamed Al Rubeai, Dr Cleo Kontoravdi and Dr Sarantos Kyriakopoulos for fruitful discussions on the work presented in this manuscript. YF was funded by Symphogen A/S and the Danish Agency for Science, Technology and Innovation under the Industrial PhD grant – Symphogen 01/2012-12/2014.

Conflict of Interest

The authors declare no financial or commercial conflict of interest.

Notation and symbols

- μ_{net}: net specific growth rate
- q_p: specific mAb secretion rate
- q_{glc}: specific glucose consumption rate
- N^c_{Xv,t_n}: number of cells accumulated up to time t_n
- N^c_{mAb,t_n}: amount of mAb accumulated up to time t_n
- N^c_{glc,t_n}: cumulative amount of glucose consumed up to time t_n
- $IVCN_{t_n}$: integral viable cell number at time t_n
- f_i: fraction of mAb glycoform i secreted over a specified time interval
- V: liquid volume in the cell culture flask
- $[X_v]$: viable cell density
- $[X_d]$: dead cell density
- $[Glc]$: glucose concentration
- $[Lac]$: lactate concentration
- $[mAb]$: antibody (product) concentration
- Q_{feed}: volumetric flowrate of feed
- Q_{out}: volumetric flowrate of sample removed from flask
- μ_g: specific growth rate
- k_d: specific death rate
- $q_{lac,secr}$: specific lactate secretion rate
- $q_{lac,cons}$: specific lactate consumption rate
- $\mu_{g,max}$: maximum specific growth rate
- $k_{d,max}$: maximum specific death rate
- α_X: Cellular carrying capacity
- $K_{m,glc}$: Monod constant for growth on glucose
- $K_{d,\mu}$: inverse saturation constant for the specific death rate
- $Y_{X/glc}$: yield coefficient of cells from glucose
- $Y_{lac/glc}$: yield coefficient of lactate from glucose consumption
- $Y_{X/lac}$: yield coefficient of cells from lactate consumption
- $K_{m,lac}$: Monod constant for lactate consumption
- $[Glyc_i]$: mAb glycoform i
- $[NS_k]$: Nucleotide sugar k (UDP-GlcNAc, GDP-Fuc, UDP-Gal or CMP-Neu5Ac)
- $[B_k]$: Nucleotide by-product k (UDP, GDP or CMP)
- V_l: linear velocity of secretory cargo through the Golgi apparatus
- z: normalized axial coordinate of the Golgi apparatus

- $\nu_{i,j}$: stoichiometric coefficient of glycoform i in reaction j
- r_j: reaction rate j
- $[E_j](z)$: concentration of enzyme j along the axial coordinate of Golgi
- $E_{j,max}$: parameter describing the peak concentration of enzyme j within Golgi
- $z_{j,max}$: parameter defining the location of the peak concentration of enzyme j within Golgi
- ω_j: parameter defining the width of the concentration profile of enzyme j
- $k_{f,j}$: turnover rate for enzyme j
- $K_{d,i/j}$: dissociation constant of mAb glycoform i from enzyme j
- $K_{d,k/j}$: dissociation constant of nucleotide sugar k from enzyme j
- $[Glyc_n]$: competing mAb glycoform substrates for the same enzyme
- $K_{d,n/j}$: dissociation constant of competing mAb glycoform n from enzyme j
- $K_{i,k/j}$: inhibition constant of by-product k on enzyme j
- $[Glyc_{(i+1)}]$: mAb glycoform with additional sugar residue added
- $K_{d,(i+1)/j}$: dissociation constant of mAb glycoform $(i+1)$ from enzyme j
- $[Man9]_{z=0}$: concentration of the Man9 mAb glycoform at the inlet of Golgi $(z=0)$
- $\overline{MW_{mAb}}$: mean mAb molecular weight
- Vol_{Golgi}: volume of the Golgi apparatus
- $Vol_{Golgi,baseline}$: volume of the Golgi apparatus at baseline productivity
- $[x_{mAb,i}]_{z=1}$: molar fraction of mAb glycoform i at the outlet of Golgi $(z=1)$
- f_{dil}: glycosylation enzyme dilution factor

References

[1] EvaluatePharma, in: Hills, P. A. (Ed.), Evaluate Ltd., London 2014, p. 38.

[2] Reichert, J. M., Antibodies to watch in 2015. *MAbs* 2015, *7*, 1-8.

[3] Jefferis, R., Glycosylation as a strategy to improve antibody-based therapeutics. *Nat. Rev. Drug. Discov.* 2009, *8*, 226-234.

[4] Chung, C. H., Mirakhur, B., Chan, E., Le, Q.-T.*, et al.*, Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N. Engl. J. Med.* 2008, *358*, 1109-1117.

[5] del Val, I. J., Kontoravdi, C., Nagy, J. M., Towards the Implementation of Quality by Design to the Production of Therapeutic Monoclonal Antibodies with Desired Glycosylation Patterns. *Biotechnol. Prog.* 2010, *26*, 1505-1527.

[6] Spahn, P. N., Lewis, N. E., Systems glycobiology for glycoengineering. *Curr. Opin. Biotechnol.* 2014, *30C*, 218-224.

[7] Jimenez del Val, I., Jedrzejewski, P. M., Exley, K., Sou, S. N.*, et al.*, in: Petrescu, S. (Ed.), *Application of Quality by Design Paradigm to the Manufacture of Protein Therapeutics*, InTech 2012.

[8] Stockmann, H., Adamczyk, B., Hayes, J., Rudd, P. M., Automated, High-Throughput IgG-Antibody Glycoprofiling Platform. *Anal. Chem.* 2013, *85*, 8841-8849.

[9] Reusch, D., Haberger, M., Kailich, T., Heidenreich, A. K.*, et al.*, High-throughput glycosylation analysis of therapeutic immunoglobulin G by capillary gel electrophoresis using a DNA analyzer. *MAbs* 2014, *6*, 185-196.

[10] Tharmalingam, T., Wu, C. H., Callahan, S., Goudar, C. T., A Framework for Real-time Glycosylation Monitoring (RT-GM) in Mammalian Cell Culture. *Biotechnol. Bioeng.* 2014.

[11] Sou, S. N., Sellick, C., Mason, A., Kyriakopoulos, S.*, et al.*, How does mild hypothermia affect monoclonal antibody glycosylation. *Biotechnol. Bioeng.* 2014, *112*, 1165-1176.

[12] Du, Z., Treiber, D., McCarter, J. D., Fomina-Yadlin, D.*, et al.*, Use of a small molecule cell cycle inhibitor to control cell growth and improve specific productivity and product quality of recombinant proteins in CHO cell cultures. *Biotechnol. Bioeng.* 2015, *112*, 141-155.

[13] Kang, S., Zhang, Z., Richardson, J., Shah, B.*, et al.*, Metabolic markers associated with high mannose glycan levels of therapeutic recombinant monoclonal antibodies. *J. Biotechnol.* 2015.

[14] Goetze, A. M., Liu, Y. D., Zhang, Z., Shah, B.*, et al.*, High-mannose glycans on the Fc region of therapeutic IgG antibodies increase serum clearance in humans. *Glycobiology* 2011, *21*, 949-959.

[15] Alessandri, L., Ouellette, D., Acquah, A., Rieser, M.*, et al.*, Increased serum clearance of oligomannose species present on a human IgG1 molecule. *Mabs* 2012, *4*, 509-520.

[16] Jefferis, R., Glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog.* 2005, *21*, 11-16.

[17] Malhotra, R., Wormald, M. R., Rudd, P. M., Fischer, P. B.*, et al.*, Glycosylation Changes of IgG Associated with Rheumatoid-Arthritis Can Activate Complement Via the Mannose-Binding Protein. *Nat. Med.* 1995, *1*, 237-243.

[18] Hodoniczky, J., Zheng, Y. Z., James, D. C., Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. *Biotechnol. Prog.* 2005, *21*, 1644-1652.

[19] Krambeck, F. J., Betenbaugh, M. J., A mathematical model of N-linked glycosylation. *Biotechnol. Bioeng.* 2005, *92*, 711-728.

[20] Hossler, P., Mulukutla, B. C., Hu, W.-S., Systems Analysis of N-Glycan Processing in Mammalian Cells. *Plos One* 2007, *2*.

[21] del Val, I. J., Nagy, J. M., Kontoravdi, C., A dynamic mathematical model for monoclonal antibody N-linked glycosylation and nucleotide sugar donor transport within a maturing Golgi apparatus. *Biotechnol. Prog.* 2011, *27*, 1730-1743.

[22] Kaveh, O., Hengameh, A., Johannes, G., Legge, M. M.*, et al., 12th IFAC Symposium on Computer Applications in Biotechnology*, IFAC, Mumbai, India 2013.

[23] Jedrzejewski, P. M., del Val, I. J., Constantinou, A., Dell, A., *et al.*, Towards controlling the glycoform: a model framework linking extracellular metabolites to antibody glycosylation. *International journal of molecular sciences* 2014, *15*, 4492-4522.

[24] Umaña, P., Bailey, J. E., A mathematical model of N-linked glycoform biosynthesis. *Biotechnol. Bioeng.* 1997, *55*, 890-908.

[25] Brandizzi, F., Barlowe, C., Organization of the ER-Golgi interface for membrane traffic control. *Nat. Rev. Mol. Cell Biol.* 2013, *14*, 382-392.

[26] Stanley, P., Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* 2011, *3*.

[27] Vallee, R. B., Williams, J. C., Varma, D., Barnhart, L. E., Dynein: An ancient motor protein involved in multiple modes of transport. *J. Neurobiol.* 2004, *58*, 189-200.

[28] Presley, J. F., Cole, N. B., Schroer, T. A., Hirschberg, K., *et al.*, ER-to-Golgi transport visualized in living cells. *Nature* 1997, *389*, 81-85.

[29] Tigges, M., Fussenegger, M., Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. *Metab. Eng.* 2006, *8*, 264-272.

[30] Guo, Y., Linstedt, A. D., COPII-Golgi protein interactions regulate COPII coat assembly and Golgi size. *J. Cell Biol.* 2006, *174*, 53-63.

[31] Fox, S. R., Patel, U. A., Yap, M. G., Wang, D. I., Maximizing interferon-gamma production by Chinese hamster ovary cells through temperature shift optimization: experimental and modeling. *Biotechnol. Bioeng.* 2004, *85*, 177-184.

[32] Trummer, E., Fauland, K., Seidinger, S., Schriebl, K., *et al.*, Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors. *Biotechnol. Bioeng.* 2006, *94*, 1033-1044.

[33] Ahn, W. S., Jeon, J.-J., Jeong, Y.-R., Lee, S. J., Yoon, S. K., Effect of culture temperature on erythropoietin production and glycosylation in a perfusion culture of recombinant CHO cells. *Biotechnol. Bioeng.* 2008, *101*, 1234-1244.

[34] Kantardjieff, A., Jacob, N. M., Yee, J. C., Epstein, E., *et al.*, Transcriptome and proteome analysis of Chinese hamster ovary cells under low temperature and butyrate treatment. *J. Biotechnol.* 2010, *145*, 143-159.

[35] Hendrick, V., Winnepenninckx, P., Abdelkafi, C., Vandeputte, O., *et al.*, Increased productivity of recombinant tissular plasminogen activator (t-PA) by butyrate and shift of temperature: a cell cycle phases analysis. *Cytotechnology* 2001, *36*, 71-83.

[36] Sung, Y. H., Song, Y. J., Lim, S. W., Chung, J. Y., Lee, G. M., Effect of sodium butyrate on the production, heterogeneity and biological activity of human thrombopoietin by recombinant Chinese hamster ovary cells. *J. Biotechnol.* 2004, *112*, 323-335.

[37] Spearman, M., Rodriguez, J., Huzel, N., Butler, M., Production and glycosylation of recombinant beta-interferon in suspension and cytopore microcarrier cultures of CHO cells. *Biotechnol. Prog.* 2005, *21*, 31-39.

[38] Santell, L., Ryll, T., Etcheverry, T., Santoris, M., *et al.*, Aberrant metabolic sialylation of recombinant proteins expressed in Chinese hamster ovary cells in high productivity cultures. *Biochem. Biophys. Res. Commun.* 1999, *258*, 132-137.

[39] Fan, Y., Del Val, I. J., Muller, C., Lund, A. M., *et al.*, A multi-pronged investigation into the effect of glucose starvation and culture duration on fed-batch CHO cell culture. *Biotechnol. Bioeng.* 2015, *112*, 2172-2184.

[40] Process Systems Enterprise, gPROMS ModelBuilder 4.1.0 (x64). 1997-2015.

[41] Shirsat, N., Mohd, A., Whelan, J., English, N. J., *et al.*, Revisiting Verhulst and Monod models: analysis of batch and fed-batch cultures. *Cytotechnology* 2015, *67*, 515-530.

[42] Linardos, T. I., Kalogerakis, N., Behie, L. A., Lamontagne, L. R., The Effect of Specific Growth-Rate and Death Rate on Monoclonal-Antibody Production in Hybridoma Chemostat Cultures. *Can. J. Chem. Eng.* 1991, *69*, 429-438.

[43] Bendiak, B., Schachter, H., Control of glycoprotein synthesis. Kinetic mechanism, substrate specificity, and inhibition characteristics of UDP-N-acetylglucosamine:alpha-D-mannoside beta 1-2 N-acetylglucosaminyltransferase II from rat liver. *J. Biol. Chem.* 1987, *262*, 5784-5790.

[44] Chen, W., Unligil, U. M., Rini, J. M., Stanley, P., Independent Lec1A CHO glycosylation mutants arise from point mutations in N-acetylglucosaminyltransferase I that reduce affinity for both substrates. Molecular consequences based on the crystal structure of GlcNAc-TI. *Biochemistry* 2001, *40*, 8765-8772.

[45] Pâquet, M. R., Narasimhan, S., Schachter, H., Moscarello, M. A., Branch specificity of purified rat liver Golgi UDP-galactose: N-acetylglucosamine beta-1,4-galactosyltransferase. Preferential transfer of of galactose on the GlcNAc beta 1,2-Man alpha 1,3-branch of a complex biantennary Asn-linked oligosaccharide. *J. Biol. Chem.* 1984, *259*, 4716-4721.

[46] Martinez-Menarguez, J. A., Geuze, H. J., Slot, J. W., Klumperman, J., Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell* 1999, *98*, 81-90.

[47] Wormald, M. R., Rudd, P. M., Harvey, D. J., Chang, S. C.*, et al.*, Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* 1997, *36*, 1370-1380.

[48] Gramer, M. J., Eckblad, J. J., Donahue, R., Brown, J.*, et al.*, Modulation of Antibody Galactosylation Through Feeding of Uridine, Manganese Chloride, and Galactose. *Biotechnol Bioeng* 2011, *108*, 1591-1602.

[49] Borys, M. C., Linzer, D. I., Papoutsakis, E. T., Ammonia affects the glycosylation patterns of recombinant mouse placental lactogen-I by chinese hamster ovary cells in a pH-dependent manner. *Biotechnol. Bioeng.* 1994, *43*, 505-514.

[50] Dohi, K., Isoyama-Tanaka, J., Tokuda, T., Fujiyama, K., Recombinant expression and characterization of N-acetylglucosaminyltransferase I derived from Nicotiana tabacum. *J. Biosci. Bioeng.* 2010, *109*, 388-391.

[51] Nolan, R. P., Lee, K., Dynamic model of CHO cell metabolism. *Metab. Eng.* 2011, *13*, 108-124.

[52] Garcia Munzer, D. G., Kostoglou, M., Georgiadis, M. C., Pistikopoulos, E. N., Mantalaris, A., Cyclin and DNA distributed cell cycle model for GS-NS0 cells. *PLoS Comput. Biol.* 2015, *11*, e1004062.

[53] Orellana, C. A., Marcellin, E., Schulz, B. L., Nouwens, A. S.*, et al.*, High-antibody-producing Chinese hamster ovary cells up-regulate intracellular protein transport and glutathione synthesis. *J. Proteome Res.* 2015, *14*, 609-618.

[54] Edros, R. Z., McDonnell, S., Al-Rubeai, M., Using Molecular Markers to Characterize Productivity in Chinese Hamster Ovary Cell Lines. *Plos One* 2013, *8*.

Figure Legends

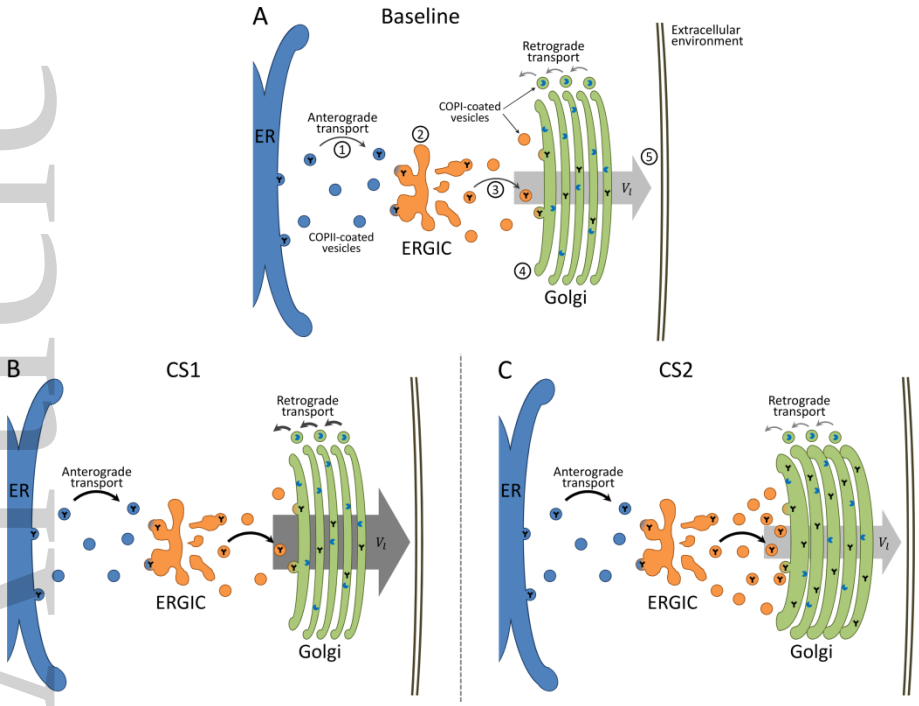

Figure 1. Mechanisms relating Golgi glycosylation with q_p [25, 26]

A) Secretory cargo ( ) is transported from the endoplasmic reticulum (ER) to the ER-Golgi intermediate compartment (ERGIC) via COPII-coated vesicles (step 1). The ERGIC concentrates the secretory cargo and aids in sorting Golgi resident proteins ( ) (step 2). The secretory cargo is then transported from ERGIC to the inlet of the Golgi space, where, if cisternal maturation is considered, the vesicles fuse to generate a new Golgi compartment (step 3). The newly formed, cargo-containing cisterna will then progress forward along the Golgi space while the protein-bound glycans are processed by the glycosylation enzymes (GEs). Upon traversing the Golgi space, the cargo is transported to the cytoplasmic membrane for secretion (step 5). This work considers two distinct mechanisms to describe changes along the secretory pathway in the event of increased specific productivity. In the first (CS1), higher production of secretory cargo is assumed to increase the linear velocity (V_l) with which the cisternae progress through the Golgi space. The rate of COPI-vesicle mediated retrograde transport is also assumed to increase so that Golgi resident proteins (nucleotide sugar transporters and GEs) are maintained at relatively constant values within Golgi (B). The second mechanism (CS2) assumes that the linear rate with which cisternae migrate through Golgi remains unchanged. Because of this, more anterograde COPII-coated vesicles accumulate at the proximal side of the Golgi space and fuse to generate larger cisternae, thus causing an overall increase in Golgi volume and a consequent dilution of resident proteins therein (C).

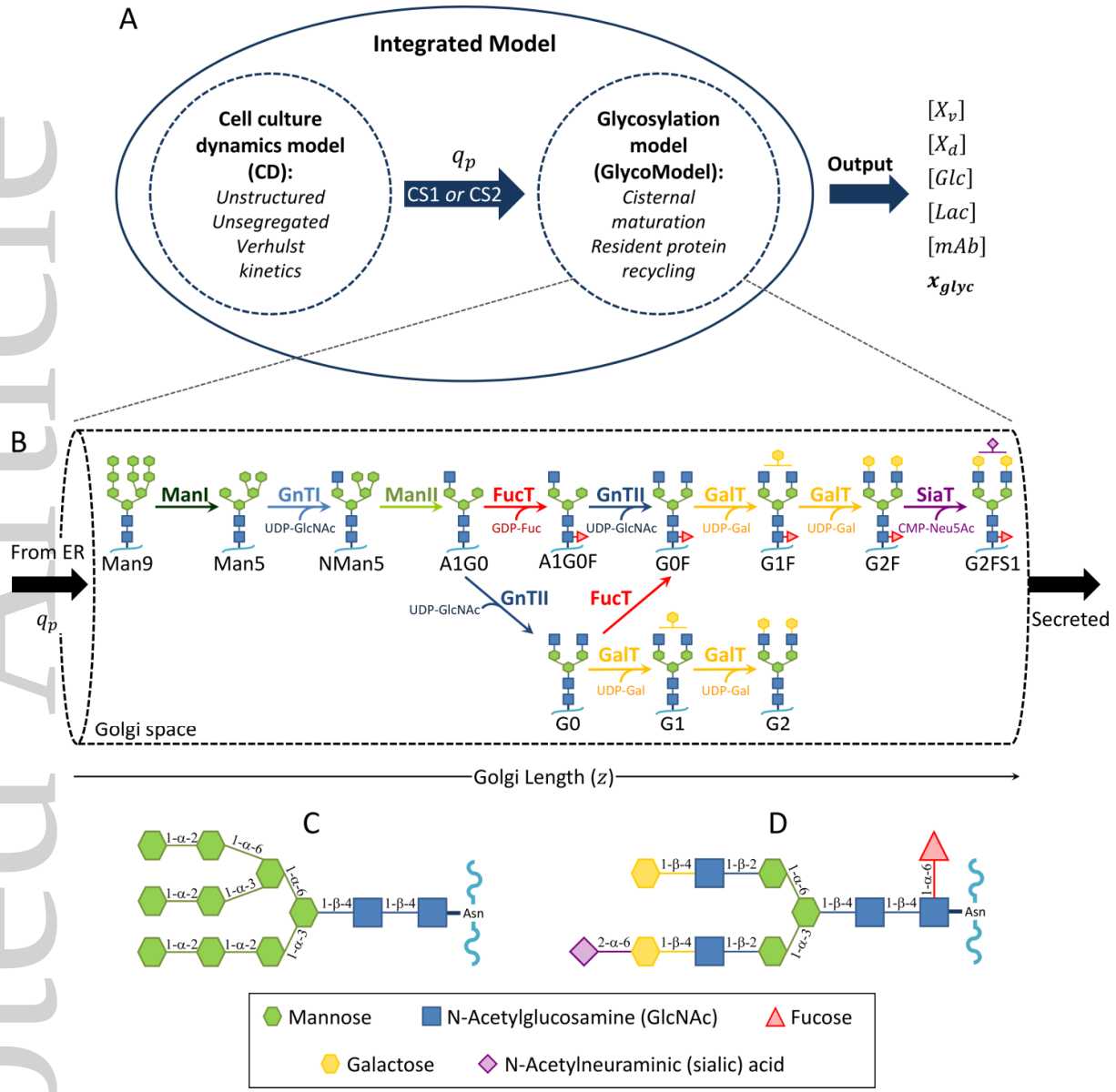

Figure 2. Model development strategy

The model development strategy employed in this work is presented (A), where an unstructured and unsegregated model for cell culture dynamics is coupled to a single-cell model for mAb Fc glycosylation via the specific productivity of the recombinant product (q_p) through the two coupling strategies (CS1 and CS2) presented in Figure 1. A simplified reaction network detailing the glycans observed in this work is also shown (B). The mAb-bound Man9 glycoform enters the Golgi apparatus (GA) from the endoplasmic reticulum (ER) and undergoes a series of enzymatically-catalysed reactions where monosaccharides are cleaved off or added to the glycan structure. Once the glycoprotein cargo traverses the GA, it is assumed to be secreted from the cell. The dashed tube surrounding the reaction network represents the Golgi space through which the cisternae carrying the secretory cargo flow [21]. Notation for all oligosaccharides is shown below the corresponding structures and the names of the enzymes

which catalyse each glycosylation reaction are presented above the arrows. The metabolic co-substrates for the monosaccharide addition reactions (nucleotide sugars) are shown below the arrows (UDP-GlcNAc, GDP-Fuc, UDP-Gal and CMP-Neu5Ac). Structures and glycosidic bond configurations for the Man9 glycan entering the GA (C) and a complex, biantennary, sialylated and fucosylated oligosaccharide structure (D) are also shown.

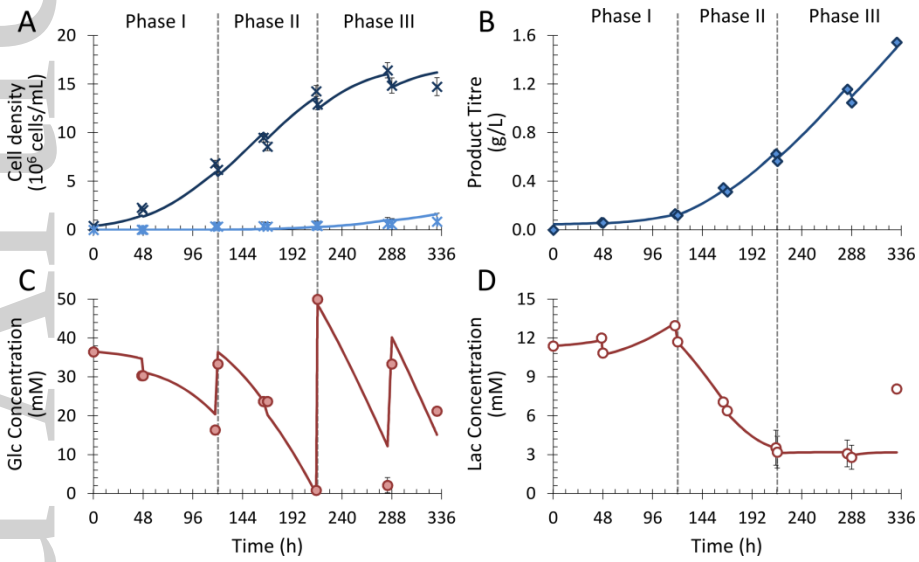

Figure 3. Cell culture dynamics

A) Experimental data for viable cell density (✕), dead cell density (✕), B) mAb titre (◆), C) extracellular glucose concentration (●) and D) extracellular lactate (○) are compared with simulation results (lines). The dashed grey lines indicate the three distinct culture phases defined in this work. Phase I is culture at 37°C, Phase II is culture at 33.5°C and Phase III is stationary growth at 33.5°C. The error bars included with the experimental data correspond to one standard deviation of duplicate experiments.

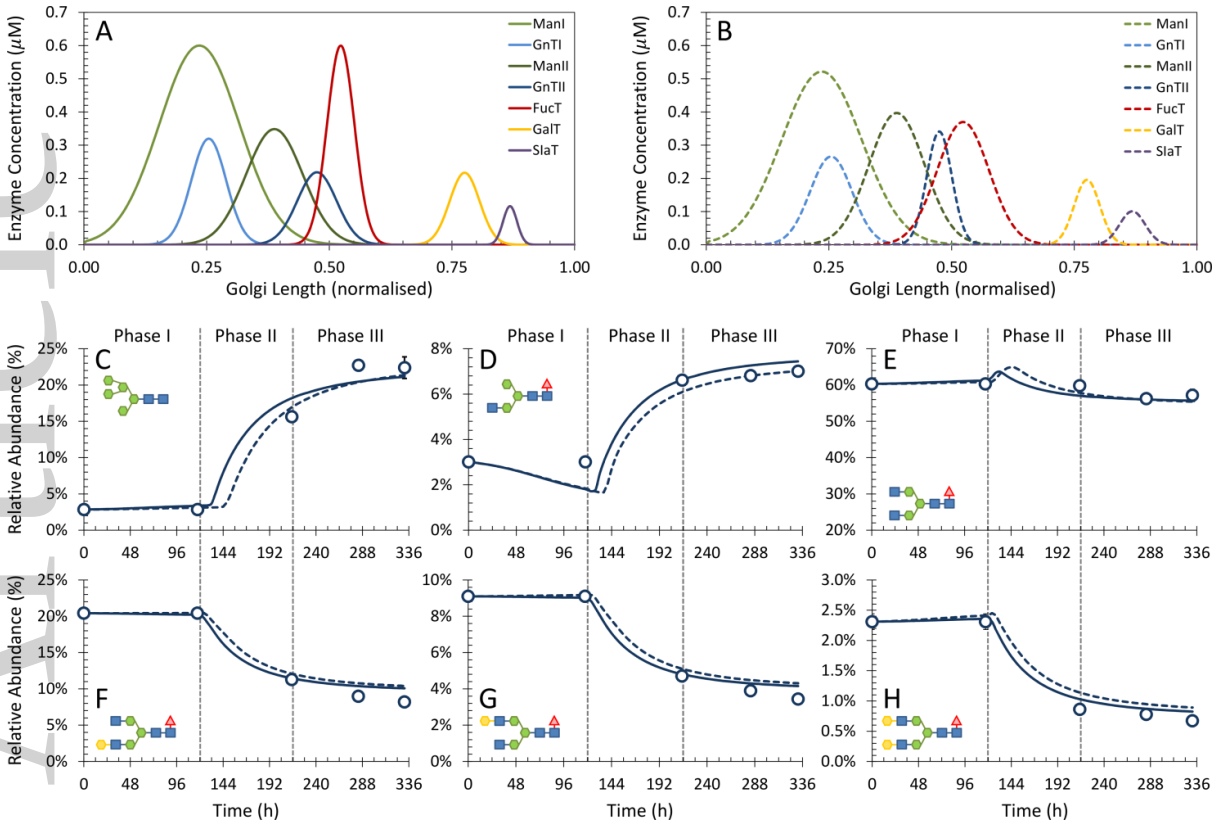

Figure 4. Comparison of model coupling strategies

The estimated distribution of GEs (ManI, GnTI, ManII, GnTII, FucT, GalT and SiaT) along the normalised length of Golgi for coupling strategies CS1 and CS2 are presented in A and B, respectively. Glycoform distribution simulation results obtained after $K_{d,i}$ estimation for CS1 and CS2 (solid and dashed lines, respectively) are compared with the experimental data (○) in C-H. Data for the six major observed mAb Fc-bound glycans is presented: Man5 (A), A1G0F (B), G0F (C), G1(1,6)F (D), G1(1,3)F (E) and G2F (F). The dashed grey lines mark the three phases of culture mentioned throughout the manuscript.

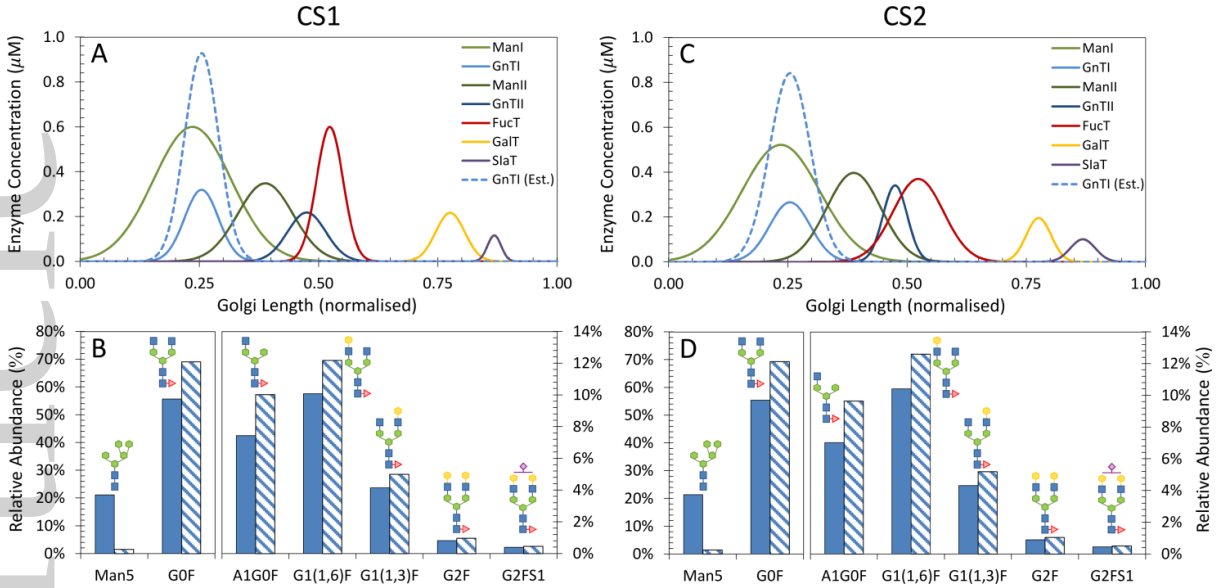

Figure 5. *In silico* minimisation of Man5 secretion

The GnTI profile that was found to minimise secretion of Man5 glycans (dashed line) is compared with the high Man5 enzyme distribution for CS1 (A). The end-point glycoform distribution calculated for CS1 with no modification to GnTI availability (solid bars) is compared with the one obtained for *in silico* GnTI overexpression (belted pattern) in (B). The same data corresponding to CS2 is presented in (C) and (D). Due to differences in scale, the axis on the left of each bar chart corresponds to the Man5 and G0F glycans and the axis on the right corresponds to the remaining five glycans.